Back-Calculating Baseline Creatinine with MDRD Misclassifies Acute Kidney Injury in the Intensive Care Unit

John W. Pickering and Zoltán H. Endre

Christchurch Kidney Research Group, Department of Medicine, University of Otago, Christchurch, New Zealand

Background and objectives: The purpose of this study was to assess the viability of back-calculation with the Modification of Diet in Renal Disease (MDRD) formula to determine baseline creatinine on the basis of acute kidney injury (AKI) metrics, RIFLE criteria, and Acute Kidney Injury Network (AKIN) criteria for the purpose of clinical trial outcomes or epidemiology.

Design, setting, participants, & measurements: This study was a retrospective analysis of prospectively collected data from patients with measured baseline creatinines before entry to the intensive care unit (ICU). The AKI status was determined using five different baseline creatinines: the measured creatinine (the standard) and an estimated creatinine determined by back-calculation using MDRD assuming a GFR of 75 ml/min ($epCr_{75}$), 100 ml/min ($epCr_{100}$), randomly generating a value on a lognormal curve ($epCr_{Rnd}$), and choosing the lowest creatinine value within the first week in the ICU ($epCr_{low}$). A subgroup of patients without chronic kidney disease (CKD) was similarly analyzed.

Results: Of 224 patients, 70 (31%) had AKI according to RIFLE and 93 (42%) according to AKIN. The $epCr_{75}$ and $epCr_{100}$ distributions greatly overestimated the proportion with AKI. The $epCr_{low}$ overestimated AKI according to AKIN but correctly estimated AKI according to RIFLE. The mean of 1000 $epCr_{Rnd}$ distributions correctly estimated AKI according to RIFLE and AKIN. Each estimated distribution performed better in the non-CKD population with the exception of $epCr_{Rnd}$. However, only the $epCr_{low}$ distribution accurately determined the proportion with AKI.

Conclusions: A measured rather than estimated value should be used for baseline creatinine in trials or epidemiologic studies of AKI.

Clin J Am Soc Nephrol 5: 1165–1173, 2010. doi: 10.2215/CJN.08531109

The consensus definitions of acute kidney injury (AKI), first by the Acute Dialysis Quality Initiative (ADQI) and later by the Acute Kidney Injury Network (AKIN), were a necessary and important step in harmonizing epidemiologic studies and clinical trials of AKI (1,2). The RIFLE (R = risk, I = injury, F = failure, L = loss, E = end-stage) and AKIN definitions stage AKI according to changes in urine output or changes in plasma creatinine from baseline. RIFLE also utilized changes in GFR as equivalent to changes in creatinine, although there was a mathematical error for R and F in this equivalence (3). The validation of these criteria has been principally through use of the plasma creatinine criteria of the first three stages of RIFLE (R, I, F) or AKIN (stages I, II, III) (Table 1). Each RIFLE and AKIN category is associated with increased mortality (4–8). The use of each classification in clinical trials and many of their strengths and limitations have been reviewed recently (9,10).

ADQI recognized that in the acute situation, measured baseline creatinines are not always available for all patients. Consequently, they recommended that an estimated baseline creatinine be calculated using the Modification of Diet in Renal Disease (MDRD) formula (11) with an assumed GFR for all patients between 75 and 100 ml/min. As a tool for *post hoc* analysis in research studies, this back-calculation method has become widespread, with most studies adopting 75 ml/min (5,8,12–22). A few epidemiologic studies have used an estimated baseline for all of their patients (*e.g.*, those using the Australian and New Zealand Intensive Care Society database (5,12–14)), whereas in other studies the proportion of patients for whom a baseline is estimated by back-calculation is as low as 7% (19).

Recent creatinine kinetic modeling has shown that using an estimated creatinine by back-calculation with the MDRD formula is likely to overestimate the proportion of patients with AKI using the RIFLE or AKIN criteria even where the estimated and measured baseline creatinine distributions have similar means (23). This observation has not been validated with clinical data. The only attempt to validate the ADQI recommendation to date observed an overestimation of the proportion of patients with AKI in a large multicenter cohort of patients already having evidence of severe AKI on admission to an intensive care unit (ICU) (24). A study of a pediatric population identified differences in AKI incidence with a baseline creatinine estimated with the Schwartz formula (25,26).

We have examined the effect on classification of AKI using an estimated rather than a measured baseline plasma creatinine

Received November 29, 2009. Accepted April 3, 2010.

Published online ahead of print. Publication date available at www.cjasn.org.

Correspondence: Dr. John Pickering, Christchurch Kidney Research Group, Department of Medicine, University of Otago–Christchurch, PO Box 4345, Christchurch 8140, New Zealand. Phone: +64-3-364-1112; Fax: +64-3-364-0935; E-mail: John.Pickering@otago.ac.nz

ISSN: 1555-9041/507–1165

Table 1. Acute kidney injury status

Criteria[a]	Creatinine-Based Definition
RIFLE	
R	pCr_{max} >1.5 × baseline plasma creatinine
I	$pCr_{max} \geq 2 \times$ baseline plasma creatinine
F	$pCr_{max} \geq 3 \times$ baseline plasma creatinine or $pCr_{max} \geq 4.0$ mg/dl with an increase ≥0.5 mg/dl above baseline plasma creatinine
Total	Sum of R, I, and F
AKIN	
I	pCr_{max} >1.5 × baseline plasma creatinine or pCr_{max} > 0.3 mg/dl above baseline plasma creatinine
II	$pCr_{max} \geq 2 \times$ baseline plasma creatinine
III	$pCr_{max} \geq 3 \times$ baseline plasma creatinine or $pCr_{max} \geq 4.0$ mg/dl with an increase ≥0.5 mg/dl above baseline plasma creatinine
Total	Sum of I, II, and III

[a]Adapted from references 1 and 2 and excluding urine output and GFR decrease criteria.

for each patient in a cohort of general ICU patients for whom baseline creatinine was available.

Materials and Methods

Population

This study included consecutive patients admitted to the ICUs of the Christchurch (for the Canterbury District Health Board, CDHB) or Dunedin (for the Otago District Health Board, ODHB) Hospitals in New Zealand between March 5, 2006 and July 8, 2008. Patients were excluded if they were younger than 16 years of age; had obvious hematuria, rhabdomyolysis and/or myoglobinuria, or polycythemia (hemoglobin >165 g/L or hematocrit >48 in women and hemoglobin >185 g/L or hematocrit >52 in men); were receiving cytotoxic chemotherapy or renal replacement therapy (RRT); or were assessed to need RRT within 48 hours or not expected to survive 72 hours. They were also excluded if they had pre-existing severe renal disease defined by plasma creatinine >3.0 mg/dl or had already experienced a greater than 3-fold rise in plasma creatinine from a known baseline or had a urine output <0.3 ml/kg per hour for >6 hours (anuric). All patients or their family gave written informed consent. The study was approved by the multiregional ethics committee of New Zealand (MEC/050020029) and registered under the Australian Clinical Trials Registry (ACTRN012606000058572 Early Acute Renal Failure 2 [EARLYARF 2], http://www.actr.org.au). Further details of the EARLYARF trial are available in Endre *et al.* (27).

For this analysis, we selected the cohort of all patients who did not receive study drug and who had available a baseline creatinine measured during an earlier admission or obtained before elective surgery.

Design

Plasma creatinine was measured on entry to the ICU, 12 and 24 hours after entry, and then daily (morning) for 7 days. Measured baseline plasma creatinine concentrations (pCr_m) were retrospectively collected from a chart review using the following rules ranked in descending order of preference: (*1*) The most recent pre-ICU value between 30 and 365 days ($n = 86$) or presurgery value for elective cardiac surgery patients at high risk of AKI ($n = 28$); (*2*) pre-ICU value >365 days if the patient age was <40 years and creatinine was stable (within 15% of the lowest ICU creatinine) ($n = 7$); (*3*) pre-ICU value >365 days if it was less than the initial creatinine on entry to ICU ($n = 58$); and (*4*) pre-ICU value 3 to 39 days if it was less than the initial creatinine on entry to ICU and not obviously AKI ($n = 45$).

Four methods were used to determine an estimated baseline creatinine for each patient: (*1*) back-calculation with the MDRD formula and assuming all patients had a normal GFR of 75 ml/min ($epCr_{75}$); (*2*) back-calculation with the MDRD formula and assuming all patients had a normal GFR of 100 ml/min ($epCr_{100}$); (*3*) randomly assigning a plasma creatinine ($epCr_{Rnd}$) along a lognormal curve with input parameters determined by fitting a lognormal curve to the measured baselines ($\mu = -0.0134$ and $\sigma = 0.3078$); and (*4*) the lowest plasma creatinine measured over 7 days from entry to the ICU ($epCr_{low}$).

Back-calculation with the MDRD formula uses the following equation:

$$epCr = \left(\frac{GFR}{Sex \times Race \times 186 \times Age^{-0.203}}\right)^{-\frac{1}{1.154}} \text{(mg/dl)}$$

where *GFR* is the assumed GFR (ml/min); $Sex = 1$ if male and 0.742 if female; $Race = 1.21$ if black, otherwise $Race = 1$; and *Age* is in years.

For each of the four methods the difference between the baseline creatinine and the maximum plasma creatinine over 7 days after entry to the ICU (pCr_{max}) was determined and used to assess the severity of AKI. AKI was determined as RIFLE status R, I, F or AKIN status I, II, or III according to Table 1. For calculation of numbers with AKI according to the $epCr_{Rnd}$ distribution, 1000 distributions were calculated and the mean number of patients in each severity category was calculated. For each of the 1000 distributions a random number generator (*lognrnd*, Matlab 2009a, MathWorks, Natick, MA, USA) distributed the patients over the same lognormal curve.

Subgroup Analysis

Patients were divided into a pre-existing chronic kidney disease (CKD) cohort (CKD determined by having an eGFR <60 ml/min calculated by MDRD using the measured baseline, pCr_m) and a non-CKD cohort.

Assays

Creatinine concentrations were determined by the Jaffe reaction using Abbott reagents on an Architect ci8000 (at the CDHB laboratory) or an Aeroset analyzer (Abbott Laboratories, Abbott Park, IL) or using Roche reagents on a Modular P Analyzer (Roche Diagnostics GmbH, Sandhofer Strasse, Mannheim, Germany) at the ODHB laboratory. Variations between laboratories in Australia and New Zealand are less than ±15% (28).

Statistical Analyses

The introduction of bias by estimating creatinine was assessed using the Bland–Altman method, which looks at the agreement between measured and estimated baseline distributions (29). The Bland–Altman method plots the mean against the difference between the measured and estimated baseline creatinine for each patient. Perfect agreement between the estimated and measured creatinines would result in all points lying along the 0 line of the *y* (difference) axis. Bias was defined as the total mean difference between estimated and measured creatinines, precision as 1 SD of the bias, and proportional bias as the slope

of the regression line of the differences between estimated and measured creatinines plotted against the average of estimated and measured creatinines. The Wilcoxon signed rank test (matching pairs, nonnormal distribution) was used to compare the entire estimated baseline creatinine distributions with the entire measured distribution, Note that this is not merely a comparison of the medians but of the entire distributions. The distribution means were compared with a paired *t* test after they were log-transformed. Correlations were nonparametric (Spearman's). Differences in proportions in the AKI classifications were compared with Fisher's exact test. The specificity and sensitivity of each estimated distribution was calculated for each individual by comparing the true RIFLE and AKIN classification (on the basis of the measured baseline, pCr_m) with that obtained using the estimated baseline creatinine. The number of true and false positives and negatives was then obtained and the sensitivity and specificity calculated.

Statistical analyses were performed with GraphPad Prism 5.0 (GraphPad Software, San Diego, CA) and Matlab 2009a (MathWorks, Natick, MA). Significance was assumed at $P < 0.05$.

Results

The EARLYARF study recruited 528 patients, of whom 84 received study drug and 220 had no available creatinine measured before entry to ICU. The remaining 224 patients comprise the cohort analyzed here. Their characteristics are summarized in Table 2. All estimated and measured creatinine distributions were approximately lognormal. Sixty-two (28%) were classified as CKD.

Only the $epCr_{100}$ distribution differed significantly from the measured distribution (pCr_m) (Wilcoxon signed rank, $P < 0.0001$, Table 3). The means of the log-transformed $epCr_{100}$ and $epCr_{low}$ values were less than that of the pCr_m distribution (paired *t* test, $P < 0.0001$ and $P < 0.001$ respectively). However, Bland–Altman plots show that there was strong proportionality bias in estimating baseline creatinine when back-calculating with MDRD (Figure 1, A and B, Table 3). The $epCr_{low}$ distribution had a small negative proportional bias and the smallest bias and greatest precision (Figure 1D, Table 3). Only the $epCr_{low}$ distribution correlated with pCr_m ($r = 0.69$, $P = 0.0013$). The distribution had the same median, but perhaps a slightly broader distribution than pCr_m (Table 3). An example of the distribution generated by randomly assigning baseline creatinines across a lognormal curve shows only a small bias but poor precision (Figure 1C).

Of the 224 patients 70 (31%) had AKI according to RIFLE and 93 (42%) had AKI according to AKIN (Table 4). All estimates of baseline creatinine led to overestimation of total AKI by RIFLE and AKIN, with the exception of $epCr_{Rnd}$ (Table 4, Figure 2). However, this difference was significant only for $epCr_{75}$ and

Table 2. Patient characteristics ($n = 224$)

Characteristic	Male ($n = 123$)	Female ($n = 101$)
Age (years)	63 ± 18	60 ± 17
Weight (kg)	84 ± 18	75 ± 22
APACHE II	19 ± 6	18 ± 6
SOFA	6.7 ± 2.8	6.5 ± 2.8
eGFR (ml/min)	80 ± 29	74 ± 27
Pre-existing CKD (eGFR < 60 ml/min)	38 (31%)	19 (19%)
Diagnostic classification		
abdominal aortic aneurysm rupture and repair	8 (7%)	0
abdominal surgery or inflammation	10 (8%)	16 (16%)
burns	2 (2%)	0
cardiac arrest	17 (14%)	10 (10%)
cardiac surgery	30 (24%)	22 (22%)
collapse, cause unknown	1 (1%)	1 (1%)
neurological surgery, injury, seizure, or hemorrhage	6 (5%)	14 (14%)
other	0	2 (2%)
pulmonary or thoracic surgery or failure	17 (14%)	17 (17%)
sepsis	23 (19%)	15 (15%)
trauma	9 (7%)	4 (4%)
Baseline pCr_m (mg/dl)		
mean ± SD	1.12 ± 0.35	0.94 ± 0.32
median (IQR)	1.02 (0.91 to 1.24)	0.86 (0.68 to 1.07)
min, max	0.57, 2.38	0.45, 2.01
ICU maximum pCr_{max} (mg/dl)		
mean ± SD	1.74 ± 1.04	1.21 ± 0.48
median (IQR)	1.47 (1.05 to 2.04)	1.13 (0.90 to 1.47)
min, max	0.69, 7.13	0.36, 3.39
Renal replacement therapy	9 (7%)	1 (1%)

Data are presented as mean ± SD, median (IQR), or *n* (%). IQR, interquartile range.

Table 3. Plasma creatinine for baseline and estimated baselines (mg/dl)

	Measured pCr_m	Estimated			
		$epCr_{75}$ (75 ml/min)	$epCr_{100}$ (100 ml/min)	$epCr_{Rand}$ (Random)[d]	$epCr_{low}$ (Lowest)
Mean ± SD (mg/dl)	1.04 ± 0.35	0.97 ± 0.14	0.75 ± 0.11[c]	1.03 ± 0.33	1.0 ± 0.42[b]
Median (IQR) (mg/dl)	0.92 (0.79 to 1.20)	1.01 (0.82 to 1.05)	0.79[c] (0.64 to 0.82)	0.99 (0.80 to 1.21)	0.91 (0.68 to 1.24)
Min, max (mg/dl)	0.45, 2.38	0.77, 1.33	0.60, 1.04	0.42, 2.33	0.23, 2.83
Correlation with pCr_m (r)	–	0.05	0.05	[e]	0.69[c]
Bias (mg/ml)		0.071	0.28	[e]	0.037
Precision (mg/ml)		0.37	0.36	[e]	0.32
Proportional bias		1.41[c]	1.60[c]	[e]	−0.22[a]

r, Spearman's ρ; Bias, total mean difference between estimated and measured creatinines; Precision, 1 SD of the bias.

Comparison of log-transformed means with pCr_m (paired t test): [b]$P < 0.001$; [c]$P < 0.0001$.

Comparison of entire distribution with the pCr_m distribution (Wilcoxson signed rank, paired): [c]$P < 0.0001$.

Proportional bias is the slope of the regression line of the differences between estimated and measured creatinines against the average of estimated and measured creatinines. A slope of 0 means no proportional bias. P value is the significance of the difference of the slope from 0: [a]$P < 0.01$, [c]$P < 0.0001$.

[d]Each statistic for $epCr_{Rand}$ is the mean of 1000 generations of 224 random creatinines across the lognormal curve.

[e]Correlation, bias, precision, and proportional bias are not given because these measurements average out to zero over 1000 generations.

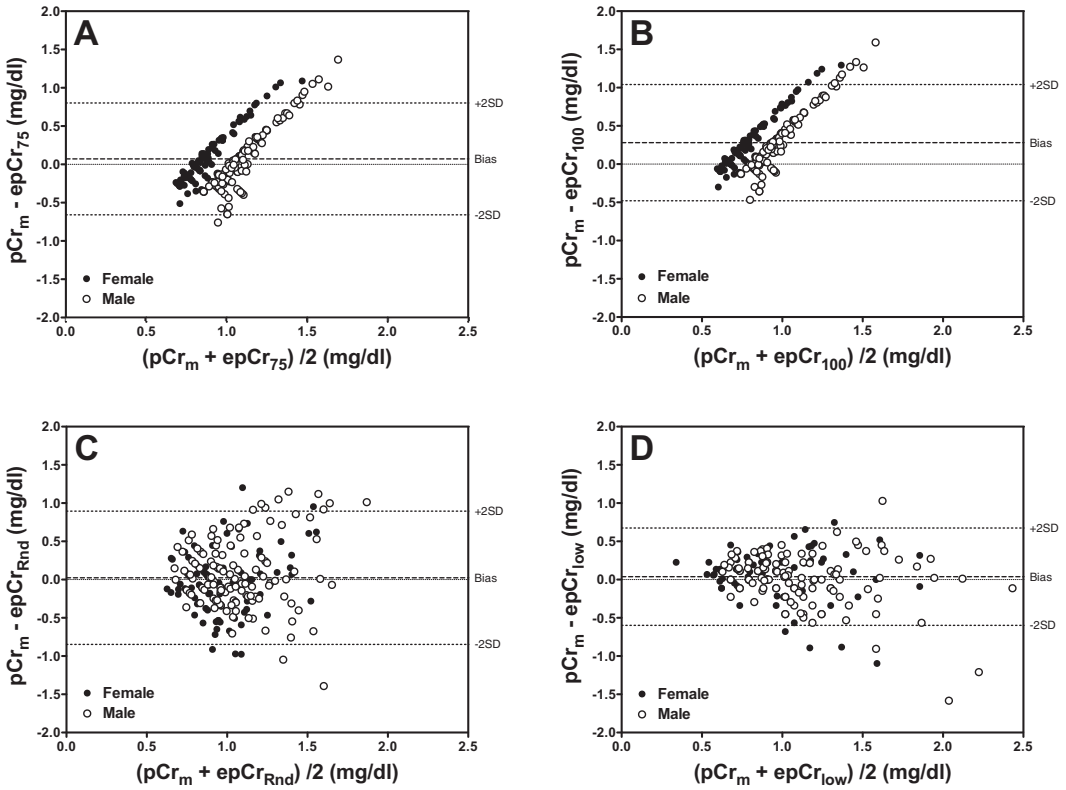

Figure 1. Bland–Altman plots of the measured creatinine (pCr_m) against the difference between the measured (pCr_m) and estimated baselines: (A) $epCr_{75}$, (B) $epCr_{100}$, (C) $epCr_{Rnd}$ (an example of the 1000 random distributions calculated), and (D) $epCr_{low}$ ($n = 224$, men $n = 123$, women $n = 101$). A perfect agreement between pCr_m and an estimated baseline distribution would mean all points lying alongside the $y = 0$ line. Total bias and SD shown are for the entire distribution (men + women).

Table 4. Number of patients at each AKI stage on the basis of RIFLE and AKIN staging ($n = 224$)

	Measured pCr_m	Estimated			
		$epCr_{75}$ (75 ml/min)	$epCr_{100}$ (100 ml/min)	$epCr_{Rand}$ (Random)[d]	$epCr_{low}$ (Lowest)
RIFLE					
R (% error)	34	47 (38.2)	53 (55.9)	36 ± 5 (5.9)	48 (41.1)
I (% error)	26	36 (38.4)	53 (104)	25 ± 4 (−3.8)	18 (−30.8)
F (% error)	10	12 (20.0)	29 (190)	7 ± 2 (−30.0)	8 (−20.0)
total (% error)	70	95 (35.7)[a]	135 (92.8)[c]	68 ± 5 (−2.9)	74 (5.7)
sensitivity (95% CI)		0.81 (0.76 to 0.86)	0.94 (0.91 to 0.97)	0.56 ± 0.05	0.60 (0.54 to 0.67)
specificity (95% CI)		0.75 (0.70 to 0.81)	0.55 (0.49 to 0.62)	0.81 ± 0.02	0.81 (0.75 to 0.86)
AKIN					
I (% error)	57	70 (22.8)	68 (19.3)	56 ± 6 (−1.8)	97 (70.2)
II (% error)	26	36 (38.5)	53 (104)	25 ± 4 (−3.8)	18 (−30.8)
III (% error)	10	12 (20.0)	29 (190)	7 ± 2 (−30.0)	8 (−20.0)
total (% error)	93	118 (26.9)[a]	150 (61.3)[c]	88 ± 5 (−5.4)	123 (32.3)[b]
sensitivity (95% CI)		0.82 (0.77 to 0.87)	0.95 (0.92 to 0.98)	0.63 ± 0.04	0.86 (0.81 to 0.91)
specificity (95% CI)		0.68 (0.62 to 0.75)	0.53 (0.47 to 0.60)	0.78 ± 0.03	0.68 (0.62 to 0.74)

CI, confidence interval; % error, the percentage difference between the number using an estimated creatinine distribution and the number using the measured baseline distribution (pCr_m).

[a]$P < 0.05$; [b]$P < 0.01$; [c]$P < 0.0001$.

[d]Each statistic for $epCr_{Rand}$ is the mean ± SD of 1000 generations of 224 random creatinines across the lognormal curve.

$epCr_{100}$. $epCr_{100}$ performed worst, overestimating AKI by 92.8% (RIFLE). AKI incidence was best estimated by $epCr_{Rnd}$ (underestimate of 2.9% for RIFLE and 5.4% for AKIN) and $epCr_{low}$ (overestimate of 5.7% on the basis of the RIFLE criteria). The specificity was higher for AKIN than for RIFLE, although this was at the expense of sensitivity (Table 4).

The $epCr_{75}$ and $epCr_{100}$ distributions overestimated all severity stages, whereas $epCr_{low}$ overestimated RIFLE R and AKIN I but underestimated the other stages (Table 4, Figure 2).

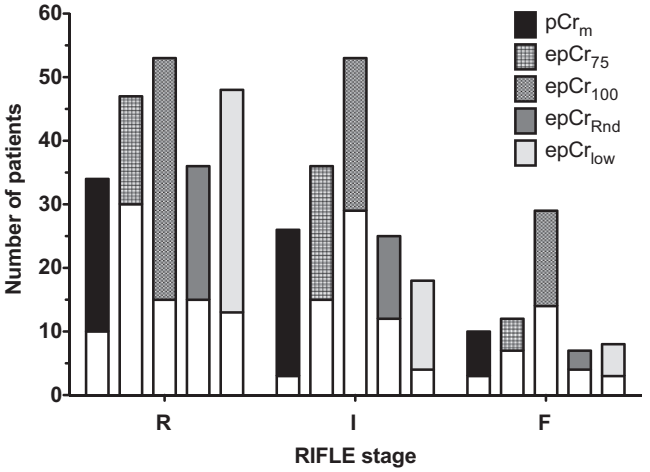

Figure 2. Dependency of patients classified as AKI using RIFLE criteria R, I, and F depending on method of selecting baseline creatinine. Each bar represents the total number of patients. The lower (unshaded) portion of the bar represents the number of CKD patients. Note that CKD patients contribute disproportionally to this overestimation.

Sensitivity Analysis

Patients with pre-existing CKD seem to account disproportionally for the overestimation at the higher (RIFLE I and F, which are equivalent to AKIN II and III) severity stages (Figure 2). Removing these patients led to underestimation with the $epCr_{75}$ and $epCr_{Rnd}$ distributions, a reduced overestimation with the $epCr_{100}$ distribution, and a good performance with the $epCr_{low}$ distribution (Table 5).

Even where total AKI incidence was correctly estimated (*e.g.*, with $epCr_{low}$, using RIFLE, in the non-CKD group) the sensitivity (0.58) and specificity (0.80) were not perfect, indicating there is still a high proportion of false negatives and false positives and the correlation with the measured baseline ($r = 0.69$) was not particularly high.

The patient cohort with most severe disease (upper quartile of Acute Physiology, Age, Chronic Health Evaluation [APACHE] scores) appeared to show less overestimation of AKI status than the cohort with the least severe disease (lower quartile of APACHE scores). Fourteen patients from the lower quartile and 37 from the upper quartile of APACHE scores had AKI according to AKIN. The $epCr_{75}$ and $epCr_{low}$ distributions overestimated total AKIN by 36% (4 patients) and 79% (11 patients) in the least-severe-disease cohort, respectively. In the most severely diseased cohort, the overestimations were 19% and 8% for the $epCr_{75}$ and $epCr_{low}$ distributions, respectively. This difference in overestimation between the least and most severely diseased cohorts was significant only for $epCr_{low}$ ($P = 0.002$).

Discussion

In this general ICU population, the use of estimated baseline creatinine distributions based on the MDRD equation led to

Table 5. Number of non-CKD patients at each AKI stage on the basis of RIFLE and AKIN staging ($n = 162$)

	Measured pCr_m	Estimated			
		$epCr_{75}$ (75 ml/min)	$epCr_{100}$ (100 ml/min)	$epCr_{Rand}$ (Random)[c]	$epCr_{low}$ (Lowest)
RIFLE					
R (% error)	24	17 (−29.2)	38 (58.3)	21 ± 4 (−12.5)	35 (45.5)
I (% error)	23	21 (−8.7)	24 (4.3)	13 ± 3 (−41.6)	14 (−39.1)
F (% error)	7	5 (−28.6)	15 (114)	3 ± 2 (−57.1)	5 (−28.6)
total (% error)	54	43 (−20.4)	77 (42.6)[a]	37 ± 4 (−31.5)[a]	54 (−0.0)
sensitivity (95% CI)		0.76 (0.69 to 0.83)	0.93 (0.89 to 0.97)	0.49 ± 0.05	0.58 (0.51 to 0.66)
specificity (95% CI)		0.98 (0.96 to 1.0)	0.75 (0.68 to 0.82)	0.90 ± 0.03	0.80 (0.73 to 0.86)
AKIN					
I (% error)	42	37 (−11.9)	52 (23.8)	31 ± 4 (−26.2)	66 (57.1)
II (% error)	23	21 (−8.7)	24 (4.3)	13 ± 3 (−41.6)	14 (−39.1)
III (% error)	7	5 (−28.6)	15 (114)	3 ± 2 (−57.1)	5 (−28.6)
total (% error)	72	63 (−12.5)	91 (26.4)[a]	48 ± 4 (−33.3)[b]	85 (18.1)
sensitivity (95% CI)		0.76 (0.70 to 0.83)	0.93 (0.89 to 0.97)	0.55 ± 0.04	0.86 (0.81 to 0.91)
specificity (95% CI)		0.91 (0.87 to 0.95)	0.73 (0.67 to 0.80)	0.90 ± 0.03	0.74 (0.68 to 0.81)

[a]$P < 0.05$; [b]$P < 0.01$.
[c]Each statistic for $epCr_{Rand}$ is the mean ± SD of 1000 generations of 224 random creatinines across the lognormal curve.

overestimation of AKI using the RIFLE or AKIN criteria. As expected, overestimation was greatest when the assumed population GFR (100 ml/min) was well above the mean (77 ± 28 ml/min) and median estimated GFR (75 [58 to 94] ml/min) values. However, even when a comparable GFR (75 ml/min) was used to back-calculate baseline creatinine, the proportion of patients was overestimated in accordance with our previous modeling (23). Although we are not aware of other equations being used to back-calculate baseline creatinine, it would be reasonable to consider other equations as alternatives to the MDRD equation in this context. We considered the Cockcroft–Gault equation (30) because it is commonly used in other contexts, and the recently introduced Chronic Kidney Disease Epidemiology Collaboration equation (31) because it is derived similarly to the MDRD equation and has been suggested as an improvement on it. Both equations yielded similar overestimations to those reported here (see online supplementary data).

In this context, back-calculating baseline creatinine with MDRD is no better than using a random number generator over a lognormal curve based on the measured baseline distribution. This is because back-calculation relies only on age and race and takes no account of actual renal function on entry to the ICU. If there is a measured baseline available for a proportion of the ICU population from which the characteristics of a lognormal equation can be deduced, then randomly fitting patients without a measured baseline to a lognormal curve on the basis of the distribution of the measured baseline values of the available cohort may be a better solution than back-calculating with MDRD. The alternative of using the lowest plasma creatinine over 7 days better approximates the true baseline distribution and leads to a similar proportion of patients being diagnosed with AKI according to the RIFLE criteria. However, the proportion diagnosed as AKI with the AKIN criteria was still overestimated, and waiting until 7 days have elapsed delays decision-making based on the estimates.

It is possible that the presence of CKD in the population is driving this overestimation. When patients with an estimated GFR <60 ml/min were removed from the population, overestimation was much reduced. Bagshaw and colleagues (24) also recently observed this phenomenon in patients in the Beginning and Ending Supportive Therapy for the Kidney (BEST Kidney) study. In that population of patients with evidence of severe AKI, back-calculation of baseline creatinine with an assumed GFR of 75 ml/min overestimated AKI (RIFLE definition) by 42%. These patients had a correlation of $r = 0.49$ between the measured and estimated baseline creatinine distributions that improved to $r = 0.90$ when the CKD patients were excluded. In our study, the corresponding correlations went from $r = 0.05$ ($P = 0.45$) to $r = 0.33$ ($P < 0.0001$) when the CKD patients were excluded. Although these correlations showed an improvement with removal of CKD patients, they remained poor. This may reflect the different populations. The inclusion criterion in the BEST study was severe AKI, as evidenced by oliguria, azotaemia, and/or the need for acute RRT. In the study presented here, patients were excluded if they were judged to need RRT within 48 hours; 73.4% of patients required RRT in the BEST study compared with 5% in our study.

Similar to the effect of baseline function, the degree of overestimation produced by back-calculation may depend on the severity of illness because overestimation was higher in populations with relatively low severity, although small numbers limit the certainty of this observation. Thus, it remains possible that MDRD extrapolated creatinine concentrations might prove useful in a selected population of patients who are sicker and have a higher incidence of severe AKI and need for RRT as in the BEST study.

Other limitations include restriction of our study population to just two centers. As evidenced by the differences between this study and the BEST study, population characteristics can have a marked effect on outcome. The use of AKIN and RIFLE in this study necessarily does not take into account the rate of increase in creatinine or the duration of the increase as in the complete AKIN and RIFLE definitions. This is inevitable when baseline measurements are many days, if not months, before entry to ICU. Thus, our conclusions are confined to *post hoc* analysis as used in clinical trials or epidemiology. They do not assist with clinical diagnosis at the time of presentation to ICU or elsewhere in patients without a baseline creatinine.

The MDRD equation was not derived from patients with normal renal function and is quite inaccurate when GFR is >60 ml/min (32). On this basis alone, the use of the equation to back-calculate baseline creatinine using GFRs of 75 to 100 ml/min should be questioned. The implication from the study presented here is that until a validated alternative to back-calculation using the MDRD formula is adopted, it should not be used for *post hoc* analysis of epidemiologic studies, clinical trials of therapeutic agents, or biomarker validation trials. The use of a measured creatinine as used here that is assumed to best represent the baseline (*e.g.*, the lowest of the ICU entry or the value after 7 days) represents a possible alternative but needs validating in other studies. Our choice may be influenced by plasma volume expansion in patients with high fluid loading. An alternative but impractical approach would be to exclude creatinine measures in the first 2 days after volume resuscitation. Another alternative would be to use a creatinine measure at, for example, 30 days post-ICU entry; however, patients developing CKD after an AKI episode may influence this. Inevitably, there will need to be compromise. However, any reasonable estimate based on creatinine measures in each individual is still likely to be better than an estimate that takes into account only age, gender, and race.

The clinical diagnosis of AKI on admission to hospital or ICU remains a problem given the absence of a baseline creatinine in approximately 50% of patients (27). Rather than "throw out the baby with the bathwater," clinicians may need to continue to use an extrapolated creatinine using the MDRD equation (with 75-ml/min GFR) until a more suitable tool is devised, or perhaps until validated biomarkers of injury highlight the possibility of AKI. Our results suggest a greater degree of confidence in using the MDRD equation when patients are more severely ill or have prior CKD.

When designing a clinical trial, a decision is required as to which measured baseline creatinine should be utilized. The most recent measurement in a non-acute situation is preferred. If this is more than approximately 12 months pre-ICU admission, the question arises whether it is a reasonable measure of the true baseline. In particular, the possibility that this is an underestimate due to developing CKD must be considered. Unfortunately, this will remain unanswerable unless the patient survives and recovers renal function to at least the same degree and thus a similar or lower plasma creatinine. Clinical trials will need to decide an *a priori* strategy to allow for this inevitable reduction in statistical power. In normal clinical use, the decision is ethically easier, because the decision is appropriately made on the best data available. However the validity of the decision is similarly only ascertainable retrospectively. This dilemma will remain until a biomarker of injury or function is available that does not require a prior baseline measurement. Acute injury biomarkers are becoming a reality for the former strategy; real-time GFR might be an option for the latter.

Unfortunately, without a valid estimate of baseline creatinine, the assessment of early biomarkers and of early intervention in AKI is presently difficult and potentially biased. Selection bias may even occur if patients have a history that includes a serum creatinine measurement because these may be more likely to have CKD, suspected kidney disease, or another chronic disease than patients presenting with no previous serum creatinine measurements. For example, young men are more likely to present with trauma in the ICU without a previous creatinine measurement than, for example, elderly patients presenting with pneumonia or a cardiac arrest.

Note that errors in baseline estimates affect patients in placebo and treatment populations in trials of AKI treatments. This means that small treatment efficacy is more likely to be missed because of "noise" in the data caused by estimating baseline creatinine. The greater the proportion of patients for whom an estimate rather than a measured baseline is used, the greater the noise. Although we have previously suggested that AKI intervention and prevention trials use a continuous measure of renal function rather than categorical measure (*e.g.*, RIFLE or AKIN) to assess efficacy (23), we highlight that continuous measures will also be influenced by errors in the estimation of baseline creatinine. In designing trials, the choice of study population should account for the fraction of patients for whom an estimated baseline will be used, even when that estimate was based on a measured creatinine during hospital stay (such as epCrlow). The larger the fraction, the larger the sample size will need to be. The alternative is to accept that an outcome based on function is inappropriate as the primary outcome and that we need to choose an outcome based on a biomarker that increases predictably with injury onset and decreases with injury cessation. Injury biomarkers are likely to be the answer, and different biomarkers may be required for onset and offset.

In conclusion, for epidemiologic studies and clinical trials, estimates of baseline creatinine that do not take into account renal function on presentation are misleading. This recommendation does not refer to management of individual patients. The dilemma of how to measure baseline renal function in patients without a prior creatinine measurement remains difficult and awaits a valid real-time estimate of GFR.

Acknowledgments

This study was supported by the Health Research Council of New Zealand grant 05/131. We thank the staff of the ICUs of Dunedin and Christchurch hospitals—John Dean, Jill Robinson, and Robyn Hutchi-

son—for assisting with data collection. We thank Associate Professor Chris Frampton for statistical support. Parts of this manuscript were presented in poster form at the American Society of Nephrology 42nd Annual Meeting and Scientific Exposition, San Diego, California, October 17 to November 1, 2009.

Disclosures

None.

References

1. Bellomo R, Ronco C, Kellum JA, Mehta RL, Palevsky P, for the Acute Dialysis Quality Initiative workgroup: Acute renal failure—Definition, outcome measures, animal models, fluid therapy and information technology needs: The Second International Consensus Conference of the Acute Dialysis Quality Initiative (ADQI) Group. *Crit Care* 8: R204–R212, 2004
2. Mehta RL, Kellum JA, Shah SV, Molitoris BA, Ronco C, Warnock DG, Levin A, for the Acute Kidney Injury Network: Acute Kidney Injury Network: Report of an initiative to improve outcomes in acute kidney injury. *Crit Care* 11: R31, 2007
3. Pickering JW, Endre ZH: GFR shot by RIFLE: Errors in staging acute kidney injury. *Lancet* 373: 1318–1319, 2009
4. Hoste EA, Clermont G, Kersten A, Venkataraman R, Angus DC, De Bacquer D, Kellum JA: RIFLE criteria for acute kidney injury are associated with hospital mortality in critically ill patients: A cohort analysis. *Crit Care* 10: R73, 2006
5. Bagshaw SM, George C, Dinu I, Bellomo R: A multi-centre evaluation of the RIFLE criteria for early acute kidney injury in critically ill patients. *Nephrol Dial Transplant* 23: 1203–1210, 2008
6. Barrantes F, Tian J, Vazquez R, Amoateng-Adjepong Y, Manthous CA: Acute kidney injury criteria predict outcomes of critically ill patients. *Crit Care Med* 36: 1397–1403, 2008
7. Ricci Z, Cruz D, Ronco C: The RIFLE criteria and mortality in acute kidney injury: A systematic review. *Kidney Int* 73: 538–546, 2008
8. Joannidis M, Metnitz B, Bauer P, Schusterschitz N, Moreno R, Druml W, Metnitz P: Acute kidney injury in critically ill patients classified by AKIN *versus* RIFLE using the SAPS 3 database. *Intens Care Med* 35: 1692–1702, 2009
9. Cruz DN, Ricci Z, Ronco C: Clinical review: RIFLE and AKIN—Time for reappraisal. *Crit Care* 13: 211, 2009
10. Endre ZH, Pickering JW: Outcome definitions in non-dialysis intervention and prevention trials in acute kidney injury (AKI). *Nephrol Dial Transpl* 25: 107–118, 2010
11. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D: A more accurate method to estimate glomerular filtration rate from serum creatinine: A new prediction equation. *Ann Intern Med* 130: 461–470, 1999
12. Bagshaw SM, George C, Bellomo R, for the ANZICS Database Management Committee: Early acute kidney injury and sepsis: A multicentre evaluation. *Crit Care* 12: R47, 2008
13. Bagshaw SM, George C, Bellomo R, for the ANZICS Database Management Committee: A comparison of the RIFLE and AKIN criteria for acute kidney injury in critically ill patients. *Nephrol Dial Transplant* 23: 1569–1574, 2008
14. Bagshaw SM, George C, Bellomo R: Changes in the incidence and outcome for early acute kidney injury in a cohort of Australian intensive care units. *Crit Care* 11: R68, 2007
15. Cruz DN, Bolgan I, Perazella MA, Bonello M, de Cal M, Corradi V, Polanco N, Ocampo C, Nalesso F, Piccinni P, Ronco C, Investigators, NEIPHROSoAKIN-A: North East Italian Prospective Hospital Renal Outcome Survey on Acute Kidney Injury (NEiPHROS-AKI): Targeting the problem with the RIFLE Criteria. *Clin J Am Soc Nephrol* 2: 418–425, 2007
16. Maccariello E, Soares M, Valente C, Nogueira L, Valença RV, Machado JE, Rocha E: RIFLE classification in patients with acute kidney injury in need of renal replacement therapy. *Intens Care Med* 33: 597–605, 2007
17. Valdivieso JRP, Bes-Rastrollo M, Monedero P, de Irala J, Lavilla FJ: Evaluation of the prognostic value of the risk, injury, failure, loss and end-stage renal failure (RIFLE) criteria for acute kidney injury. *Nephrology* 13: 361–366, 2008
18. Perez-Valdivieso JR, Bes-Rastrollo M, Monedero P, de Irala J, Lavilla FJ: Prognosis and serum creatinine levels in acute renal failure at the time of nephrology consultation: An observational cohort study. *BMC Nephrology* 8: 14, 2007
19. Ahlstrom A, Kuitunen A, Peltonen S, Hynninen M, Tallgren M, Aaltonen J, Pettila V: Comparison of 2 acute renal failure severity scores to general scoring systems in the critically ill. *Am J Kidney Dis,* 48: 262–268, 2006
20. Uchino S, Bellomo R, Goldsmith D, Bates S, Ronco C: An assessment of the RIFLE criteria for acute renal failure in hospitalized patients. *Crit Care Med* 34: 1913–1917, 2006
21. Constantin J-M, Futier E, Perbet S, Roszyk L, Lautrette A, Gillart T, Guerin R, Jabaudon M, Souweine B, Bazin J-E, Sapin, V: Plasma neutrophil gelatinase-associated lipocalin is an early marker of acute kidney injury in adult critically ill patients: A prospective study. *J Crit Care* 25: 176.e1–176.e6, 2010
22. Udy A, O'Donoghue S, D'Intini V, Healy H, Lipman J: Point of care measurement of plasma creatinine in critically ill patients with acute kidney injury. *Anaesthesia* 64: 403–407, 2009
23. Pickering JW, Frampton CM, Endre ZH: Evaluation of trial outcomes in acute kidney injury by creatinine modeling. *Clin J Am Soc Nephrol* 4: 1705–1715, 2009
24. Bagshaw SM, Uchino S, Cruz D, Bellomo R, Morimatsu H, Morgera S, Schetz M, Tan I, Bouman C, Macedo E, Gibney N, Tolwani A, Oudemans-van Straaten H, Ronco C, Kellum JAM: A comparison of observed *versus* estimated baseline creatinine for determination of RIFLE class in patients with acute kidney injury. *Nephrol Dial Transpl* 24: 2739–2751, 2009
25. Schwartz GJ, Haycock GB, Edelmann CM, Spitzer A: Simple estimate of glomerular-filtration rate in children derived from body length and plasma creatinine *Pediatrics* 58: 259–263, 1976
26. Zappitelli M, Parikh CR, Akcan-Arikan A, Washburn KK, Moffett BS, Goldstein SL: Ascertainment and epidemiology of acute kidney injury varies with definition interpretation. *Clin J Am Soc Nephrol* 3: 948–954, 2008
27. Endre ZH, Walker RJ, Pickering JW, Shaw GM, Frampton CM, Henderson SJ, Hutchison R, Mehrtens J, Robinson JM, Schollum JBW, Westhuyzen J, Celi LA, MacGinley R, Campbell IJ, George, PM: Early intervention with erythro-

poietin does not affect the outcome of acute kidney injury (the EARLYARF trial). *Kidney Int* 2010 [Epub ahead of print, doi: 10.1038/ki.2010.25]
28. Jones GRD, Mathew T, Johnson D, Peake M: Implementation of the routine reporting of eGFR in Australia and New Zealand. *Scand J Clin Lab Inv* 68: 23–29, 2008
29. Bland J, Altman D: Statistical-methods for assessing agreement between 2 methods of clinical measurement. *Lancet* 327: 307–310, 1986
30. Cockcroft D, Gault M: Prediction of creatinine clearance from serum creatinine. *Nephron* 16: 31–41, 1976
31. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J: A new equation to estimate glomerular filtration rate. *Ann Intern Med* 150: 604–613, 2009
32. Coresh J, Auguste P: Reliability of GFR formulas based on serum creatinine, with special reference to the MDRD study equation. *Scand J Clin Lab Inv* 68: 30–38, 2008